

doi:10.19306/j.cnki.2095-8110.2025.01.009

顾及 GNSS 水汽线性和非线性特征的 ERA5 PWV 产品校正方法

郭宏武^{1,2}, 李浩杰³, 杜芬玲⁴, 李祖锋⁵, 马智³,
赵庆志³, 翟园¹, 杜玉柱⁶, 岳延兵⁶

(1. 西安市气象局, 西安 710032; 2. 南京气象科技创新研究院中国气象局交通气象重点开放实验室, 南京 210041; 3. 西安科技大学测绘科学与技术学院, 西安 710054;
4. 咸阳市勘察测绘院, 陕西咸阳 712000; 5. 中国电建集团西北勘测设计研究院有限公司, 西安 710065;
6. 山西水利职业技术学院, 山西运城 044000)

摘要:高精度大气可降水量(PWV)对数值天气预报和短临极端天气研究等具有重要意义。第五代欧洲中尺度天气预报中心再分析数据集(ERA5)能够提供高时空分辨率的 PWV 产品,但在局部区域其精度并不理想,无法满足区域精细化天气预报预警的现实需求。为提高 ERA5 PWV 产品的局部区域精度,提出了一种顾及 GNSS 水汽线性和非线性特征的 ERA5 PWV 校正方法。该方法考虑了 GNSS PWV 与 ERA5 PWV 之间的系统偏差,利用 Lomb-Scargle 周期图方法分析了 PWV 偏差周期项,基于最小二乘原理建立了 PWV 偏差的线性周期校正模型。其次,综合考虑了线性校正后 PWV、不同气象参数和时空因子对水汽残差的影响,基于反向传播神经网络(BP-NN)构建了不同季节 ERA5 PWV 非线性校正模型,以优化 ERA5 PWV 局部区域精度。选取中国大陆构造环境监测网络 2021—2023 年 GNSS, ERA5 和无线电探空站的 PWV,以及气象站实测数据进行实验。结果表明,该方法在不同时空水汽对比上均具有较好的精度,与 ERA5 PWV 产品相比其均方根(RMS)平均改善率为 32.15%。该方法能够有效改善局地 ERA5 PWV 精度,为区域精细化天气预警预报研究等提供高精度的水汽信息。

关键词:GNSS; ERA5 PWV; Lomb-Scargle 周期图; 反向传播神经网络

中图分类号:P228.4

文献标志码:A

文章编号:2095-8110(2025)01-0097-14

The ERA5 PWV correction method considering linear and nonlinear characteristics of GNSS PWV

GUO Hongwu^{1,2}, LI Haojie³, DU Fenling⁴, LI Zufeng⁵, MA Zhi³, ZHAO Qingzhi³,
ZHAI Yuan¹, DU Yuzhu⁶, YUE Yanbing⁶

(1. Xi'an Meteorological Bureau, Xi'an 710032, China; 2. Key Laboratory of Transportation Meteorology of China Meteorological Administration, Nanjing Joint Institute for Atmospheric Sciences, Nanjing 210041, China;
3. College of Geomatics, Xi'an University of Science and Technology, Xi'an 710054, China;
4. Xianyang Institute of Surveying and Mapping, Xianyang, Shaanxi 712000, China;

收稿日期: 2024-07-12; 修訂日期: 2024-10-25

基金项目:国家自然科学基金(42274039);陕西省气象局秦岭和黄土高原生态环境气象重点实验室开放研究基金(2023K-1);山西省水利科学技术研究推广项目(2024GM10, 2024GM11);南京气象科技创新研究院北极阁开放研究基金(BJG202411)

作者简介:郭宏武(1969—),男,高级工程师,主要从事大气水汽反演及其创新应用方面的研究。

通信作者:赵庆志(1989—),男,博士,教授,主要从事 GNSS 数据处理及其气象创新应用方面的研究。

-
5. Power China Northwest Engineering Corporation Limited, Xi'an 710065, China;
 6. Shanxi Conservancy Technical Institute, Yuncheng, Shanxi 044000, China)

Abstract: Highly accurate precipitable water vapor (PWV) is of great importance for numerical weather prediction, short-range extreme weather prediction studies, etc. The fifth generation European centre for medium-range weather forecasts reanalysis (ERA5) is capable of providing PWV products with high spatiotemporal resolution. However, its accuracy is not ideal in local regions and cannot meet the practical needs for regionally refined weather forecasts and warnings. In order to effectively improve the accuracy of ERA5 PWV products in the local region, an ERA5 PWV correction method considering linear and nonlinear characteristics of GNSS PWV is proposed. This method considers the systematic deviation between GNSS PWV and ERA5 PWV, analyses the period term of PWV deviation by using the Lomb-Scargle (LS) periodogram method, and develops a linear period correction model for PWV deviation based on the least squares principle. In addition, the influence of linearly corrected PWVs, various meteorological parameters and spatiotemporal factors on the PWV residuals is comprehensively considered. The nonlinear correction model of ERA5 PWV in different seasons is developed based on the back propagation neural network (BP-NN), and the accuracy of ERA5 PWV in local regions is optimized. The PWV derived from the crustal movement observation network of China (CMONOC) GNSS stations, ERA5, and radiosondes stations, as well as the meteorological data measured by meteorological stations from 2021 to 2023 are selected for the experiment. The results show that the proposed method has good accuracy in different spatiotemporal water vapor contrasts, and its root mean square(RMS) average improvement rate is 32.15% compared with the ERA5 PWV product. The proposed method can effectively improve the accuracy of local ERA5 PWV and provide high-precision water vapor information for regional refined weather warning and forecasting research.

Key words: Global navigation satellite system (GNSS); The fifth generation European centre for medium-range weather forecasts reanalysis (ERA5) precipitable water vapor (PWV); Lomb-Scargle periodogram; Back propagation neural network (BP-NN)

0 引言

大气水汽作为陆面-大气能量交换的媒介,具有显著的时空变异性,对天气预报、全球水文循环及气候变化等研究具有重要影响。大气水汽通常通过大气可降水量(precipitable water vapor, PWV)量化表示。因此,获取高精度和高时空分辨率的局地 PWV,对提升天气、气候和环境监测能力,以及人们深入理解地球气候系统至关重要。

现有 PWV 获取方法主要包括无线电探空仪(radiosonde, RS)、遥感卫星、数值天气预报(numerical weather prediction, NWP)和全球卫星导航系统(global navigation satellite system, GNSS)等^[1-3]。其中,利用 RS 提供的实测数据可获取高精度 PWV(RS PWV),常被认为是校准和验证其他水汽获取方法的参考^[4]。但由于 RS PWV 存在站点

稀疏、成本高和易受极端天气影响等缺陷,在短临天气预警预报等领域的应用有限^[5-6]。遥感卫星反演 PWV 具有全天候、大范围和高空间分辨率等优势,但遥感数据易受到云与天气的影响,导致大气水汽反演精度较低^[7],且其在局部地区精度有限^[8]。NWP 模型,如欧洲中期天气预报中心(European Centre for Medium-range Weather Forecasts, ECMWF)的第五代大气再分析数据集(the fifth generation ECMWF reanalysis, ERA5),具有较高的精度和时空分辨率^[9],但在一些高 PWV 量级的局部区域其精度相对较低。GNSS 水汽反演技术具有高精度、高时空分辨率和低成本等优势,能够很好地弥补传统 PWV 探测手段的不足^[7,10]。虽然联合多源水汽能够弥补单一水汽探测技术的不足,但多种水汽探测技术获取的 PWV 间存在时空分辨率不一致、精度不等和

系统偏差等问题^[11]。因此,众多学者致力于利用高精度GNSS PWV对高时空分辨率的再分析资料获取的PWV进行校正,以获取兼具高时空分辨率和高精度的PWV产品。

现有PWV校正方法主要分为线性校正和非线性校正两种。其中,线性校正方法通过建立多源PWV间的线性模型,对低质量PWV进行改正。文献[12-13]以高精度GNSS PWV为参考,评估了不同区域中分辨率成像光谱仪(moderate resolution imaging spectroradiometer, MODIS) PWV精度,并利用GNSS PWV建立了线性MODIS PWV校正模型,提高了MODIS PWV在不同区域的精度。文献[14]利用高精度的GNSS PWV辅助风云三号A星L1级数据,建立自适应季节校正模型,反演得到高精度PWV格网产品,相较于风云三号A星L2级产品,其改善率为74.5%。文献[15]基于MODIS与ERA5数据,提出了一种增强时空自适应数据融合方法,利用青藏高原地区季风时期的数据生成高精度PWV产品,发现该产品的均方根(root mean square, RMS)介于1.1~2.0 mm之间。然而,上述研究多是基于高精度GNSS PWV构建的线性模型对PWV进行校正,难以捕捉PWV与其他气象因子及时空因子间的非线性变化。基于此,大量学者利用神经网络强大的非线性拟合能力,构建PWV非线性校正模型,以获取高精度PWV数据^[16]。文献[17]提出了一种基于反向传播神经网络(back propagation neural network, BP-NN)的MODIS PWV反演算法,基于地表覆盖类型和时空因子实现高精度PWV的获取,与MODIS PWV产品相比,反演的PWV精度提高了37.98%~68.67%。文献[18]基于广义回归神经网络,利用高精度的GNSS PWV对低质量的MODIS PWV进行优化和校正,发现RMS改善率可达37.1%。文献[19]通过分析GNSS PWV与ERA5 PWV的精度,并综合考虑ERA5 PWV和时空因子与GNSS PWV之间的关系,建立了具有附加约束条件的BP-NN融合校正模型,该模型获取的PWV的RMS改善率为18.54%。上述基于神经网络的PWV校正模型,以原始PWV和时空因子为输入变量,以高精度PWV为输出变量,虽然在一定程度上能够捕捉PWV和时空因子的非线性变化,但却忽略了气象因子对PWV的影响。同时,上述基于神经网络的PWV非线性校正模型常忽略输入变量与输出变量间的线性关系,在一定程度上影响了PWV神经网络校正模型的训练

效率和精度。

鉴于此,本文针对ERA5水汽在局部区域精度不高的现状,提出了一种顾及GNSS水汽线性和非线性特征的ERA5 PWV产品校正方法。该方法首先建立了GNSS与ERA5 PWV偏差的线性校正模型,并引入Lomb-Scargle周期图(Lomb-Scargle periodogram, LS)方法确定PWV偏差的周期项,从而实现ERA5 PWV的线性校正。此外,以线性校正后的PWV、不同气象参数和时空因子为输入,分季节建立了基于GNSS水汽的非线性校正模型,进一步实现ERA5 PWV的局地非线性精化。与ERA5 PWV产品数据相比,校正后的ERA5 PWV平均改善率达到32.15%,可为局部区域数值预报或极端天气预警等研究提供高精度的PWV产品。

1 研究区域与数据介绍

1.1 研究区域

研究区域位于中国山西省及其周边地区(北纬109°~115°,东经34°~41°),该区域地形自西北向东南倾斜,东南部多为起伏的丘陵与沟壑,而西北地区地势较为平缓^[20]。该地区的气候条件与地理特征复杂,水土流失与缺水问题严重^[21]。本文选择该区域作为实验区域,并采用2021—2023年研究区域内的ERA5格网点数据、GNSS站数据(49个)及GNSS站址的气象站数据进行实验。此外,选取无线电探空站数据以验证本文提出方法的可靠性。图1所示为研究区位置及站点分布,实验数据详细信息如表1所示。

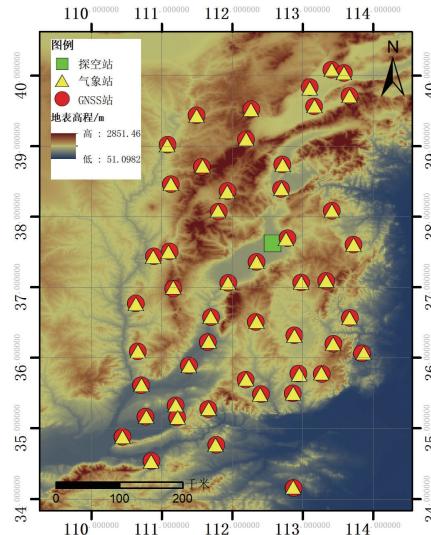


图1 研究区及GNSS、气象和无线电探空站点分布

Fig. 1 Geographical distributions of GNSS, meteorological and radiosonde stations in the experimental region

表 1 实验数据的具体信息
Tab. 1 Specific information on experimental data

数据	时间分辨率/h	空间分辨率	数据来源
GNSS	1	station	https://data.cosmic.ucar.edu/gnss-ro
RS	12	station	https://www.ncdc.noaa.gov/pub
气象数据	1	station	http://data.cma.cn/
ERA5	1	0.25°×0.25°	https://cds.climate.copernicus.eu

1.2 数据与方法

1.2.1 GNSS PWV 数据

中国大陆构造环境监测网络(crustal movement observation network of China, CMONOC)始建于1997年,以超过260个连续基准站为主体,提供的数据具有稳定性好、精度高和范围广等特点,被广泛应用于地球科学等研究领域^[22]。本文选用山西及其周边地区CMONOC和连续运行参考站共49个GNSS站2021—2023年的数据进行实验。其中天顶对流层延迟(zenith tropospheric delay, ZTD)可通过精密单点定位技术解算获取^[14],其时间分辨率为1 h。天顶对流层干延迟(zenith hydrostatic delay, ZHD)可根据测站气压和经验模型(如Saastamoinen模型)计算得出。GNSS PWV计算所需的相关气象参数由ERA5提供^[14]。首先,利用式(1)获取ZHD,通过ZTD减去ZHD获取高精度的天顶对流层湿延迟(zenith wet delay, ZWD)。然后将ZWD利用式(2)转化为PWV。

$$Z_{\text{HD}} = \frac{0.002\ 276\ 8P}{1 - 0.002\ 66\cos(2\varphi) - 0.000\ 28H} \quad (1)$$

$$P_{\text{wv}} = \frac{10^6}{(k'_2 + k_3/T_m) \cdot R_w \cdot \rho_w} \times Z_{\text{wd}} \quad (2)$$

式中,P为地表气压,单位为mbar(1 mbar=0.1 kPa);H为测站大地高,单位为km;φ为站点对应纬度,单位为(°)。 ρ_w 表示液态水的密度(1 000 kg/m³), R_w 为水蒸气比气体常量(461 J/(kg·K)), k'_2 与 k_3 为大气折射率常数,其值分别为(17±10) K/hPa和(3.776±0.04)×10⁵ K²/hPa, T_m 可通过经验模型获取,PWV详细计算步骤可参照文献[14]。

1.2.2 RS PWV 数据

美国国家海洋和大气管理局的国家气候数据中心发布了全球综合无线电探空仪资料(integrated global radiosonde archive, IGRA),该数据集包含准确的气象数据,如压强、温度和相对湿度等信息,时间分辨率为12 h^[23]。利用无线电探空数据对地

面到探测剖面顶部的比湿进行积分可获取PWV,具体计算方法如式(3)所示。本文选择与GNSS并址的RS站以验证模型校正后结果的准确性。

$$P_{\text{wv}} = \frac{1}{g} \int_{P_0}^{P_t} q dp \quad (3)$$

式中,q为比湿,单位为g/kg;g为重力加速度,单位为cm/s²;P₀是地表压力,单位为mbar;P_t是大气顶部压力,单位为mbar。

1.2.3 ERA5 PWV 数据

ERA5是ECMWF发布的第五代再分析资料。ERA5以综合预测系统(integrated forecasting system, IFS)Cy41r2为基础,利用四维变分(4D-Var)同化方案,将不同观测系统测量结果集成到大气模型中,提供地表至80 km高度内的137个垂直层上的大气参量数据,该数据集的最高空间分辨率为0.25°×0.25°,时间分辨率为1 h^[9-10]。ERA5可提供全球范围内的PWV数据,具有较高的时空分辨率,时间覆盖范围自1940年至今,常被用于研究全球和局地水汽变化,为气候变化研究提供了全面的数据支持^[6,24]。本文选取研究区域内2021—2023年的ERA5 PWV数据作为模型校正的输入数据。为获取GNSS PWV位置处的ERA5 PWV,首先基于黄良珂等^[25]的PWV垂直改正模型,将GNSS测站周围4个格网位置的ERA5 PWV沿垂直方向改正到GNSS站高度。之后,基于双线性插值方法,将4个格网位置处高程改正后的ERA5 PWV插值到GNSS站点,进而获取GNSS站点位置的ERA5 PWV。

1.2.4 气象数据

研究中使用的气象数据来自中国气象局(China meteorological administration, CMA)。CMA提供的气象数据包括气温、气压、相对湿度、降水及风速等,其时间分辨率为1 h。本文选取研究区域2021—2023年49个气象站的气温、气压、相对湿度和降水数据进行实验。为保证气象数据的质量,对气象数据进行质量控制,剔除大于3倍中误差的异常气象数据。

1.2.5 BP-NN 理论

BP-NN模型最早由Rumelhart等^[26]在1986年提出,是一种广泛应用的前馈神经网络,可用于处理回归和分类问题^[27]。BP-NN具有强大的非线性拟合能力,其整体框架如图2所示。BP-NN的计算过程包含正反向两个阶段,其中输入层数据先通过正向计算传播到隐藏层,隐藏层数据通过激活函数传播到输出层。当输出层的实际输出值与期望输出值误差不满足要求时,BP-NN模型根据输出结果与期望结果误差的大小从后向前进行反向计算,不断地修正不同神经元之间的权值和阈值,使误差呈负梯度下降到最小,最终获取最优的输出结果。

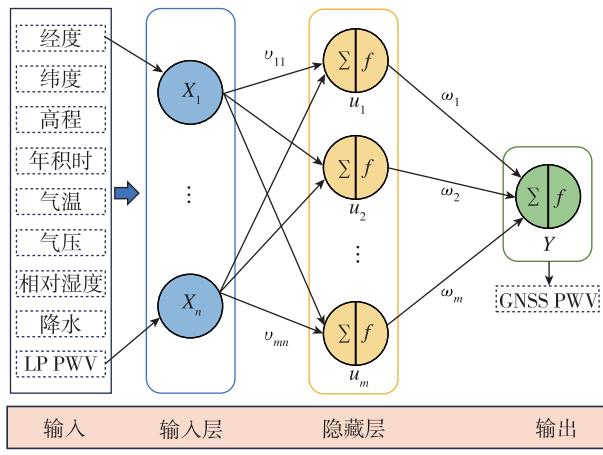


图2 BP-NN模型的框架

Fig. 2 Framework of the BP-NN model

BP-NN通常包括输入层、隐藏层和输出层。输入层负责接收外部输入数据,隐藏层是网络中的中间层,负责对输入数据进行特征提取和抽象,输出层用于输出最终的预测结果。BP-NN是将一组输入变量通过非线性的方法映射到一组输出变量中,输入层与隐藏层之间的数学关系可用式(4)表达。

$$u_j = f(\sum_{i=1}^n v_{ij} x_i + \theta_j) \quad (4)$$

式中, u_j 表示神经元节点中的映射; n 表示输入层神经元个数; v_{ij} 表示输入层与隐藏层之间的权重; x_i 为输入变量; θ_j 表示输入层与隐藏层之间的误差。此外,隐藏层与输出层之间的数学关系为

$$y = f(\sum_{j=1}^m \omega_j u_j + \theta_y) \quad (5)$$

式中, y 表示神经网络预测的输出结果; m 表示隐藏层神经元节点个数; ω_j 表示隐藏层与输出层之间的权重; θ_y 表示输出误差。在所有神经元中均存在一个映射(激活函数)将不同层串联起来,也就是上述

式中的 f 。最后,神经网络利用损失函数判断模型预测结果与真实结果之间的差距,并根据损失函数是否收敛来判断模型是否终止迭代,损失函数的具体表达式如下

$$J = \frac{1}{2} \sum_{k=1}^N (\hat{y}_k - y_k)^2 \quad (6)$$

式中, J 表示损失函数; k 为迭代次数; N 为选取迭代次数的最大值; \hat{y}_k 为输出层的真值; y_k 为神经网络模型预测值。

2 顾及GNSS水汽线性和非线性特征的ERA5 PWV校正方法

针对现有ERA5 PWV在局部区域精度较低的现状,提出了一种顾及GNSS水汽线性和非线性特征的ERA5 PWV校正方法。首先,对相关数据进行预处理,包括PWV数据反演、气象数据获取及异常数据剔除等;其次,计算GNSS站点上GNSS与ERA5水汽的时序偏差,利用LS方法确定水汽偏差周期项,并建立基于水汽偏差的线性周期校正模型,实现ERA5 PWV的线性校正;最后,以时空因子、气象因子和线性校正后的ERA5 PWV作为输入,构建基于BP-NN的PWV非线性校正模型,实现顾及水汽局地非线性特征的ERA5 PWV产品精化。

2.1 数据预处理

1)并址站选取。选用与GNSS站并址的气象站数据作为GNSS站的气象数据。并址站选取原则为经纬度差值小于0.1°且高度差小于100 m^[28]。基于上述原则,在研究区域共选取49组并址站。

2)ERA5 PWV与站点匹配。由于ERA5 PWV为格网数据,其与GNSS PWV在空间上并不匹配。因此,利用双线性插值方法获取GNSS站位置处的ERA5 PWV数据^[29],具体计算公式如下

$$\begin{aligned} p_s &= p_{a1} \cdot \frac{x_2 - x}{x_2 - x_1} \cdot \frac{y - y_1}{y_2 - y_1} + \\ &\quad p_{a2} \cdot \frac{x - x_1}{x_2 - x_1} \cdot \frac{y - y_1}{y_2 - y_1} + \\ &\quad p_{a3} \cdot \frac{x_2 - x}{x_2 - x_1} \cdot \frac{y_2 - y}{y_2 - y_1} + \\ &\quad p_{a4} \cdot \frac{x - x_1}{x_2 - x_1} \cdot \frac{y_2 - y}{y_2 - y_1} \end{aligned} \quad (7)$$

式中, p_s 为GNSS站位置处的ERA5 PWV; $p_{a1} \sim p_{a4}$ 为距离GNSS站位置最近的4个格网的ERA5 PWV,它们对应的坐标分别为 $a_1(x_1, y_1), a_2(x_2, y_2), a_3(x_1, y_1)$ 和 $a_4(x_2, y_1)$ 。

2.2 GNSS 辅助 ERA5 的 PWV 残差线性周期校正

2.2.1 GNSS PWV 和 ERA5 PWV 精度分析

本文利用 2021—2023 年的 RS PWV 分别评估了 GNSS PWV 与 ERA5 PWV 的精度。首先, 基于平均值法将 GNSS PWV 和 ERA5 PWV 的时间分辨率聚合到与 RS PWV 相一致(12 h), 进而分别获取 RS PWV 与 GNSS PWV 和 ERA5 PWV 之间的偏差。图 3 所示为以 2021—2023 年 RS PWV 为参考, GNSS PWV 和 ERA5 PWV 与 RS PWV 的长时序偏差及散点密度图。由图 3 可知, 相较于 ERA5 PWV 与 RS PWV 之间的偏差, GNSS PWV 与 RS PWV 之间的偏差更小。RS PWV 与 GNSS PWV 和 ERA5 PWV 的整体 RMS/Bias 分别为 2.02 mm/1.36 mm 和 5.22 mm/3.19 mm。结果表明, 以 RS PWV 为参考, 相较于 ERA5 PWV, GNSS PWV 精度更高, 可作为校正 ERA5 PWV 的数据源。

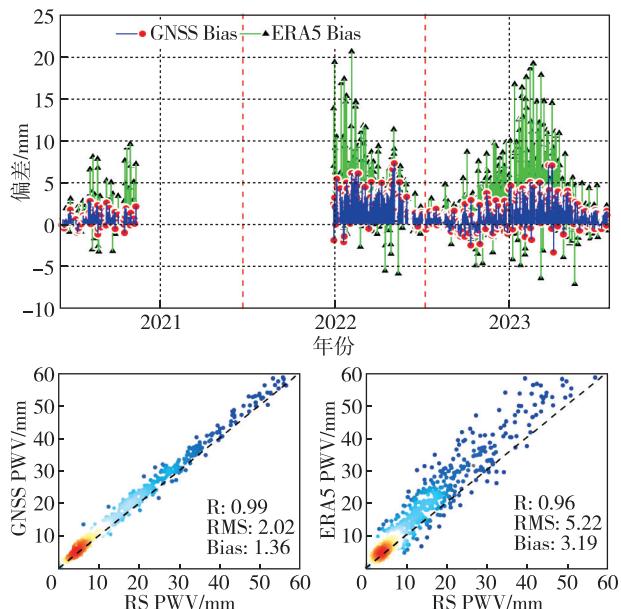


图 3 2021—2023 年 RS PWV 与 GNSS PWV 和 ERA5 PWV 的长时序偏差及散点密度图

Fig. 3 Long-term deviation and scatter density plots comparing GNSS PWV and ERA5 PWV against RS PWV from 2021 to 2023

2.2.2 PWV 残差周期项确定

为确定站点上 PWV 残差长时序的周期性, 利用 LS 方法探测非均匀采样数据的周期信号。对 49 个并址站上的 GNSS 与 ERA5 的 PWV 偏差进行周期项检验, 图 4 所示为 GNSS 站(53666 站)的 PWV

偏差长时序变化及其 LS 方法结果。由图 4 可以看出, 2021—2023 年的 PWV 偏差存在明显的年与半年周期性变化现象。图 4 中 LS 的显著性检验结果表明, PWV 偏差存在显著的年和半年周期。因此, 本文构建了基于年和半年周期的 PWV 残差线性周期模型。

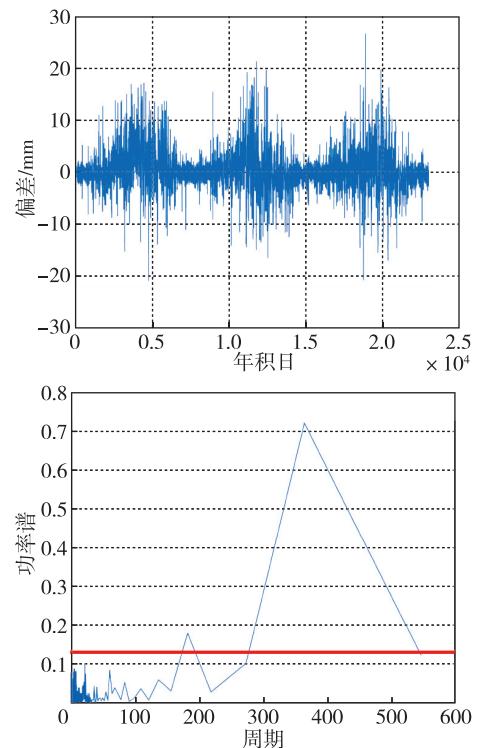


图 4 2021—2023 年站点 ERA5 与 GNSS PWV 偏差周期分析

Fig. 4 Periodic analysis of the deviation between ERA5 and GNSS PWV at a certain station from 2021 to 2023

2.2.3 水汽线性周期校正模型构建

基于确定的 PWV 残差年周期与半年周期, 构建了 PWV 残差线性校正模型^[22], 该模型具体表达式如下

$$P_{\text{WVB}} = A_0 + A_1 \cos\left(\frac{D_{\text{doy}}}{365.25} 2\pi\right) + B_1 \sin\left(\frac{D_{\text{doy}}}{365.25} 2\pi\right) + A_2 \cos\left(\frac{D_{\text{doy}}}{365.25} 4\pi\right) + B_2 \sin\left(\frac{D_{\text{doy}}}{365.25} 4\pi\right) \quad (8)$$

式中, P_{WVB} 为 GNSS PWV 和 ERA5 PWV 差值; A_0, A_1, A_2, B_1 和 B_2 为模型系数; D_{doy} 表示年积日(day of year, DOY)。

2.2.4 ERA5 PWV 线性校正

利用 GNSS PWV 和 ERA5 PWV 获取 GNSS 站点位置的 PWV 残差序列, 基于最小二乘原理获

取站点位置处的周期项系数,进而构建以GNSS站点为基准的水汽线性周期校正模型,如式(8)所示。由于最邻近法能够表达站点附近变量的空间特征^[30-32],因此,选取该方法将站点位置处获取的周期项系数扩展到ERA5对应的格网点上,进而利用式(9)对格网点上的ERA5 PWV进行线性校正^[24]。

$$P_{\text{WVLP}} = P_{\text{WVera5}} - P_{\text{WVBias}} \quad (9)$$

式中, P_{WVLP} 为 ERA5 PWV 线性校正后结果,为表述方便,本研究将其简称为 LP PWV。

2.3 基于BP-NN的ERA5 PWV残差非线性校正

大气水汽除受季节等线性因素影响外,还受时间、位置及气象等非线性因素影响^[33-34]。为进一步考虑水汽非线性的影响,本节引入BP-NN构建水汽残差非线性校正模型,对区域ERA5 PWV进行精化。

2.3.1 ERA5水汽非线性校正模型构建

ERA5水汽非线性校正模型将站点经纬度、高程、年积日、气温、气压、相对湿度、降水以及LP PWV作为输入参数,利用高精度的GNSS PWV作为非线性校正后的ERA5 PWV真值,在模型训练时将其作为输出参数,进而构建基于BP-NN的ERA5水汽非线性校正模型。该模型在使用时需输入经纬度、高程、年积日、气温、气压、相对湿度、降水以及LP PWV,即可获得非线性校正后的ERA5 PWV(本研究将其简称为LNL PWV),模型具体表达式如下

$$P_{\text{WV(GNSS PWV)}} = f(L_{\text{on}}, L_{\text{at}}, H, D_{\text{doy}}, P, T, H_{\text{um}}, P_{\text{re}}) + P_{\text{WVLP}} \quad (10)$$

2.3.2 BP-NN关键参量确定

隐藏层数和隐藏层神经元个数是影响BP-NN模型预测准确性的关键因素。其中,隐藏层的数量需要根据具体问题和数据特点来确定。由于大多数有界连续函数通常使用单隐藏层神经网络逼近任意小的误差,因此,本研究在BP-NN模型中只选用一个隐藏层^[35]。隐藏层神经元个数是影响BP-NN模型表现的另一个重要参数,不合理的神经元个数设置可能会导致模型出现不拟合或过拟合现象,进而影响模型的计算效率和精度^[36]。本文基于后验方法确定不同季节下的最优隐藏层神经元个数。根据经验研究,神

经元个数一般介于 $2(X)^{0.5} + Y$ 到 $2X + 1$ 之间, X 代表输入层节点数, Y 代表输出层节点数,因此,文中 X 与 Y 分别为 9 和 1^[17]。

基于此,从不同季节中随机选取 70% 的经纬度、高程、年积时日、气温、气压、相对湿度、降水、LP PWV 和 GNSS PWV 数据作为训练集,利用训练集数据以经纬度、高程、年积时日、气温、气压、相对湿度、降水和 LP PWV 为输入数据,以 GNSS PWV 为输出数据,建立 ERA5 水汽非线性校正模型,并将剩余的 30% 数据作为验证集对模型精度进行验证^[37]。设置隐藏层神经元个数为 7~19,利用训练集数据基于不同神经元个数构建不同季节的BP-NN 模型,并利用验证集获取上述不同 BP-NN 模型的 RMS,如图 5 所示。由图 5 可知,4 个季节的最优隐藏层神经元个数分别为 15, 17, 18 和 16。因此,选用不同季节对应的最优隐藏层神经元个数,建立不同季节的BP-NN 模型,进而获取非线性校正后的 ERA5 PWV(即 LNL PWV)。此外,选取 Relu 函数作为隐藏层与输出层的激活函数。

3 验证与评估

3.1 整体精度评估

利用验证集数据计算 LP PWV 与 LNL PWV 的内符合精度。图 6 所示为 49 个 GNSS 站不同季节下 GNSS PWV 与 ERA5 PWV、LP PWV 和 LNL PWV 对比的散点密度图。由图 6 可以看出,LP PWV 相比 ERA5 PWV 在不同季节的精度均有改善。同时,本文提出的 BP-NN 模型获得的 LNL PWV 具有最优的内符合精度,其平均 RMS 为 2.85 mm,夏季 RMS 相对较大,冬季 RMS 相对较小,这主要是由于不同季节 PWV 量级差异导致^[14]。表 2 所示为 LP PWV 和 LNL PWV 相较于 ERA5 PWV 在不同季节的 RMS 改善率。由表 2 可以看出,与 ERA5 PWV 相比,LNL PWV 在夏季改善效果最好,改善率高达 39.6%。此外,在不同季节 LP PWV 和 LNL PWV 的平均改善率分别为 24.13% 与 32.15%,证明了本文提出的 BP-NN 校正模型对 ERA5 PWV 精度具有较大改善。

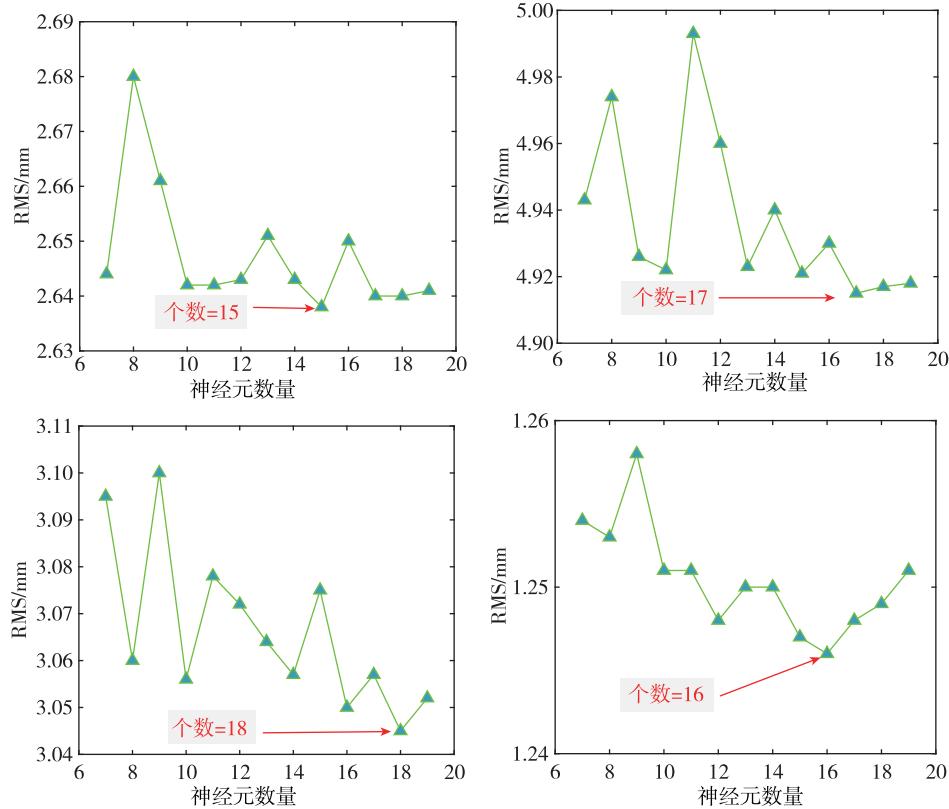


图 5 基于不同神经元数的 BP-NN 模型的 RMS 比较结果

Fig. 5 The RMS comparison results of BP-NN models based on different numbers of neurons

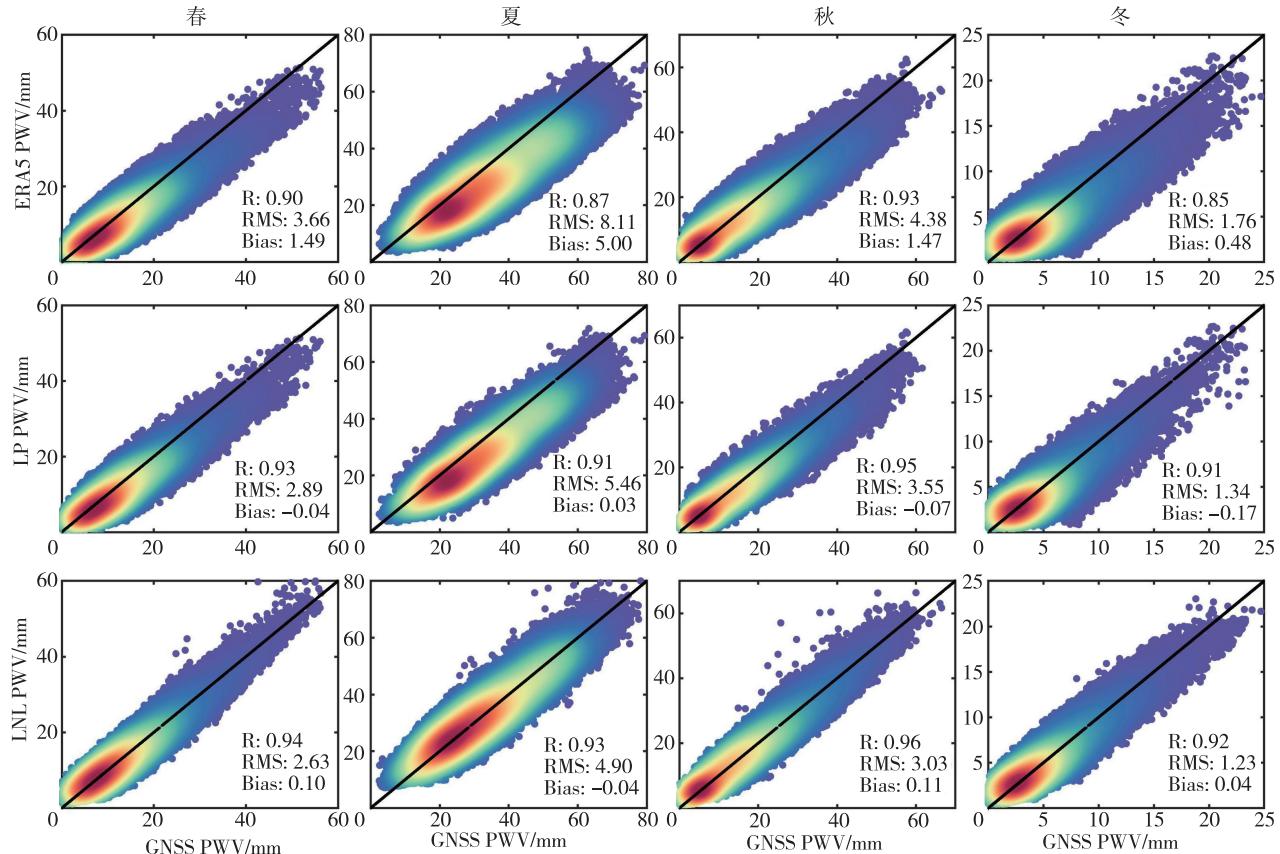


图 6 不同季节 GNSS PWV 与 ERA5/LP/LNL PWV 的散点密度对比结果

Fig. 6 Comparison of scatter density of GNSS PWV and ERA5/LP/LNL PWV in different seasons

表2 LP PWV和LNL PWV与ERA5

PWV相比在不同季节的精度改善率

Tab. 2 Accuracy improvement rate of LP PWV and LNL PWV in different seasons compared with ERA5 PWV

数据类型	RMS 改善率 %				
	春	夏	秋	冬	平均
LP VS ERA5	21.00	32.70	18.90	23.90	24.13
LNL VS ERA5	28.10	39.60	30.80	30.10	32.15

为了进一步验证 LP PWV 与 LNL PWV 的外符合精度,以 2021—2023 年研究区域内与 GNSS 站并

址的 RS PWV 为参考,计算了不同季节下 ERA5 PWV,GNSS PWV,LP PWV 和 LNL PWV 的精度。图 7 所示为 RS PWV 与 ERA5 PWV,GNSS PWV,LP PWV 和 LNL PWV 的 MAE 和 RMS 统计结果。由图 7 可以看出,ERA5 PWV,GNSS PWV,LP PWV 和 LNL PWV 在春、秋和冬季的 RMS 均小于 5 mm,夏季的 MAE 与 RMS 偏大,这与上述 PWV 量级的影响一致。相较于 ERA5 PWV 和 LP PWV,LNL PWV 在 4 个季节的精度均最优,说明了本文提出的顾及水汽线性和非线性特征的 ERA5 PWV 校正方法具有较好的改善效果和稳定性。

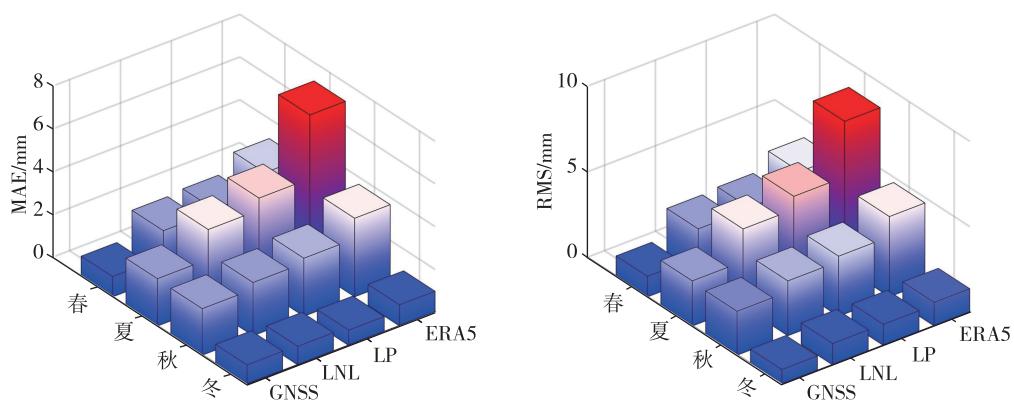


图 7 2021—2023 年 RS PWV 与 ERA5/GNSS/LP/LNL PWV 在不同季节的 MAE 和 RMS

Fig. 7 The MAE and RMS between RS PWV and ERA5/GNSS/LP/LNL PWV in different seasons from 2021 to 2023

3.2 长时序 PWV 精度评估

为了验证 LP PWV 与 LNL PWV 在不同历元的表现情况,对 2023 年 49 个站点对应的 ERA5 PWV,LP PWV 与 LNL PWV 进行计算,并求取 PWV 的日均值。以 2023 年 GNSS PWV 日均值为参考,图 8 所示为 GNSS/RS PWV 分别与 ERA5/LP/LNL PWV 长时序和偏差对比结果。由图 8 可知,LP/LNL PWV 与 GNSS/RS PWV 均具有很好的一致性。统计发现,分别以 GNSS/RS PWV 为参考,ERA5 PWV,LP PWV 与 LNL PWV 的平均偏差分别为 2.76 mm/3.07 mm,0.64 mm/0.86 mm 和 0.52 mm/0.78 mm。此外,在夏季与秋季,ERA5/LP PWV 与 GNSS/RS PWV 存在较大偏差;而 LNL PWV 在夏秋季节与 GNSS/RS PWV 的偏差依旧较小,进一步说明了 LNL PWV 在不同时序下均具有较好的鲁棒性。

3.3 PWV 空间特性评估

为了分析本文提出方法校正后 ERA5 PWV 的

空间表现能力,以 2021—2023 年 49 个 GNSS 站的训练集 PWV 为参考,评估了 ERA5 PWV,LP PWV 和 LNL PWV 在 4 个季节下的站点精度,如图 9 所示。由图 9 可知,相较于 GNSS PWV,LNL PWV 精度在不同季节均最优,LP PWV 次之,ERA5 PWV 最差。此外,由图 9 和图 1 可以发现,ERA5 PWV 的 RMS 偏大的站点均分布在地势较低的区域,其原因可能是 PWV 含量变化和高度密切相关^[29]。为了进一步验证站点精度与 PWV 量级之间的关系,图 10 所示为 2021—2023 年 49 个 GNSS 站点上不同方法获取的 PWV 均值分布。由图 10 可以看出,不同季节的 PWV 量级不同,夏季最高、秋季次之、春季和冬季最小,进一步佐证了本文提出方法在夏季对 ERA5 PWV 校正效果最优的原因。表 3 所示为 49 个站点不同 PWV 数据在 4 个季节下平均 PWV 与 RMS 统计情况。由表 3 可知,LNL PWV 与 GNSS PWV 最为接近,且在 4 个季节空间上精度最高,其平均 RMS 分别为 2.62 mm,4.91 mm,3.06 mm 与 1.24 mm。相较于 GNSS PWV,ERA5

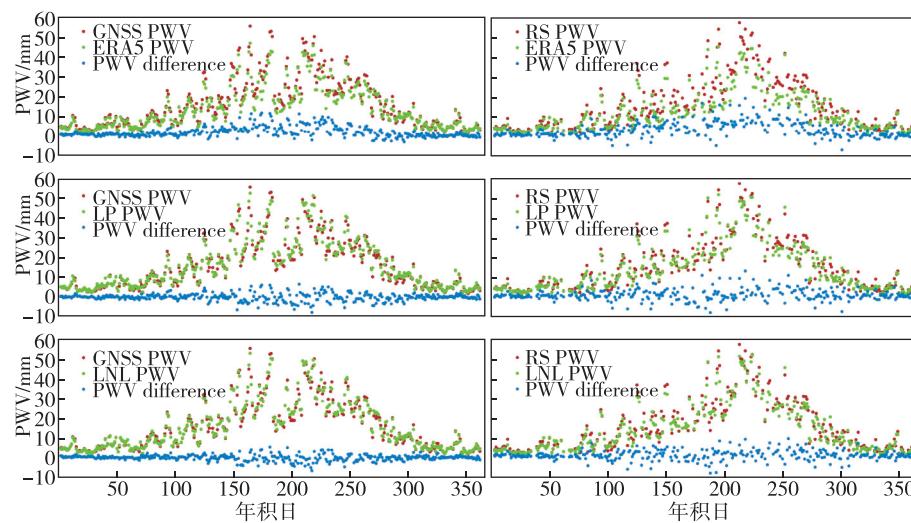


图 8 2023 年的 ERA5/LP/LNL PWV 及其分别与 GNSS/RS PWV 偏差随时间序列变化结果

Fig. 8 ERA5/LP/LNL PWV and its deviation from GNSS/RS PWV over time series in 2023

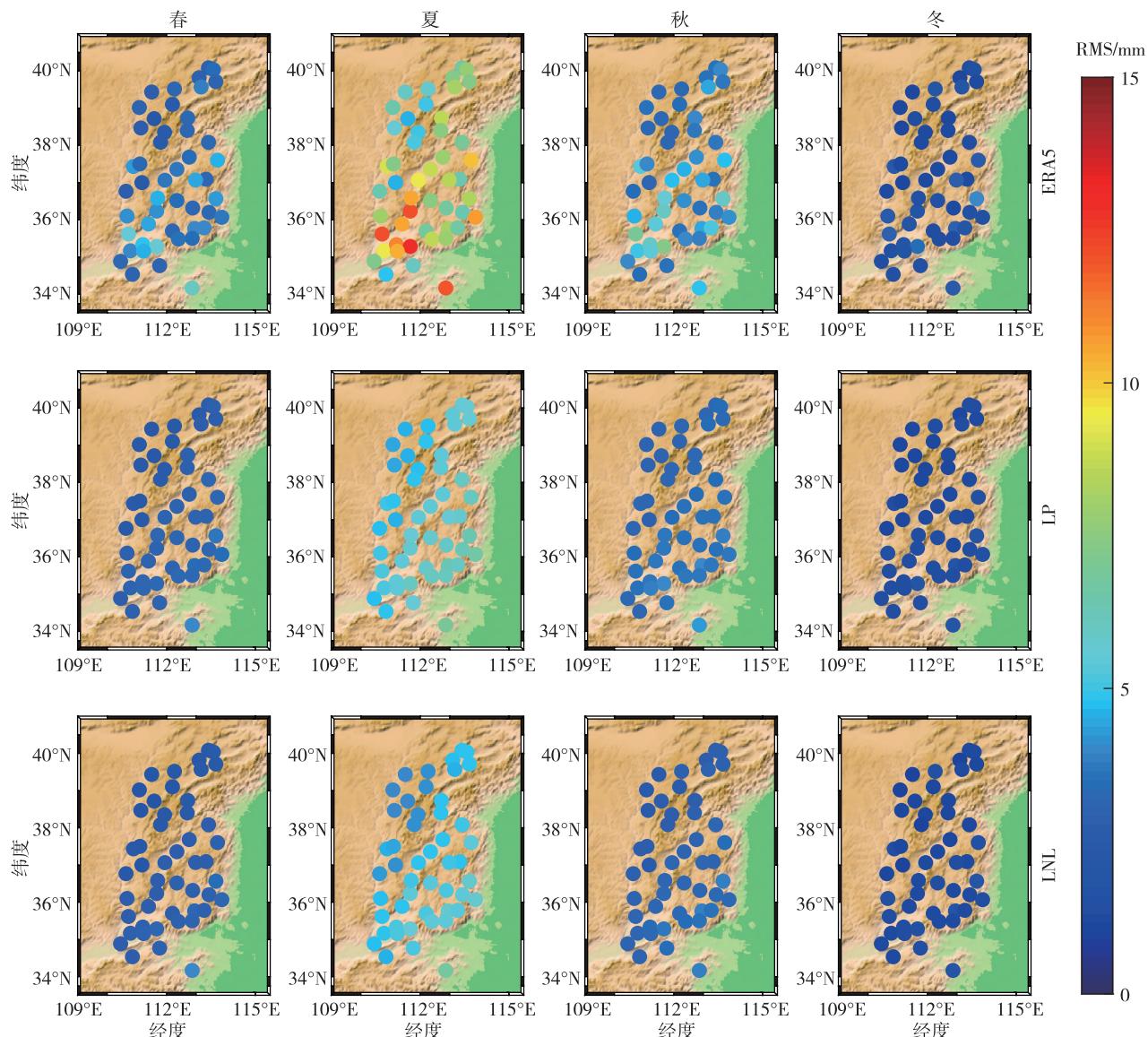


图 9 2021—2023 年不同季节的 GNSS PWV 与 ERA5/LP/LNL PWV 在不同站点位置的 RMS 分布结果

Fig. 9 The RMS distribution results between GNSS PWV and ERA5/LP/LNL PWV at different station locations for different seasons from 2021 to 2023

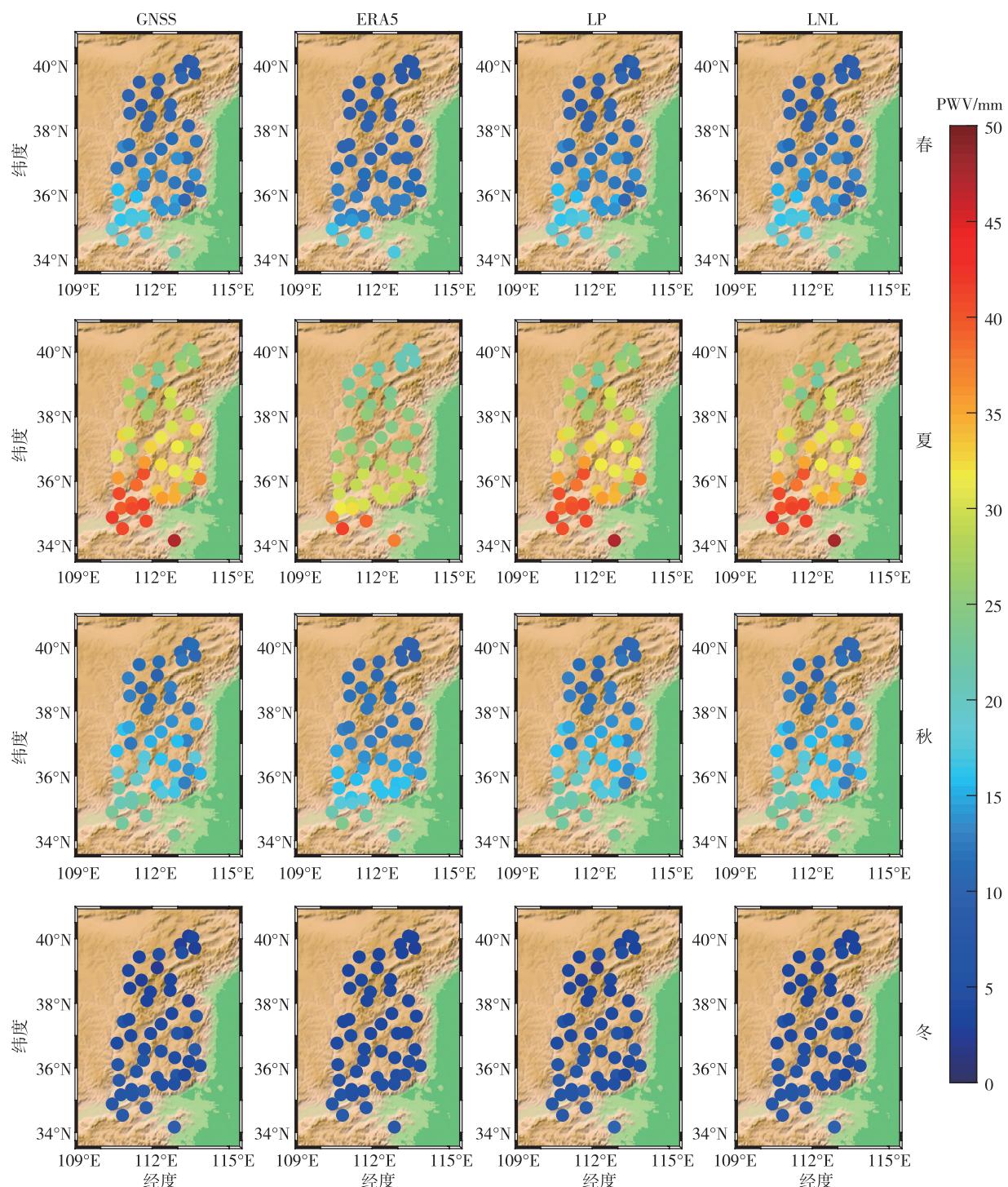


图 10 不同站点 GNSS/ERA5/LP/LNL PWV 在 2021~2023 年的平均值分布结果

Fig. 10 Distribution results of the average values of GNSS/ERA5/LP/LNL PWV at different stations from 2021 to 2023

表3 研究区域内站点的平均 PWV 与 RMS 统计结果

Tab. 3 Statistical results of average PWV and RMS for all stations in the experimental region mm

数据类型		统计结果				
		春	夏	秋	冬	平均
PWV	GNSS	12.63	32.16	15.63	4.79	16.30
	ERA5	11.12	27.06	14.13	4.28	14.15
	LP	12.74	32.09	15.74	4.99	16.39
RMS	LNL	12.59	32.19	15.61	4.76	16.29
	ERA5 VS GNSS	3.59	7.92	4.28	1.68	4.37
	LP VS GNSS	2.88	5.45	3.44	1.34	3.28
	LNL VS GNSS	2.62	4.91	3.06	1.24	2.95

PWV 存在明显的低估现象。此外,在空间上 LNL PWV 的平均 RMS 为 2.95 mm,相比 ERA5 的精度 (RMS 为 4.37 mm)有了很大的提升,进一步说明了本文所提方法具有较好的稳健性与鲁棒性。

4 结论

针对现有 ERA5 PWV 在局部区域精度低的缺陷,提出了一种顾及 GNSS 水汽线性和非线性特征的 ERA5 PWV 校正方法。该方法利用 LS 方法分析 PWV 偏差周期项,基于最小二乘原理建立 PWV 偏差的线性周期校正模型。联合时空因子、气象因子及 LP PWV,建立基于 BP-NN 的 ERA5 PWV 非线性校正模型,以克服非线性因素变化对 PWV 校正的影响。通过对分析新方法获取的 PWV 与 GNSS PWV 和 ERA5 PWV 之间的精度,得出以下结论:

1)相较于 GNSS PWV,新模型获取的 LNL PWV 精度得到较大的改善,其平均 RMS 为 2.85 mm。与 ERA5 PWV 相比,LNL PWV 在夏季改善效果最好,改善率高达 39.6%,LP/LNL PWV 的 RMS 平均改善率分别为 24.13%/32.15%。以 2021—2023 年 GNSS 站址的 RS PWV 为参考,外符合精度验证结果也证实了 LNL PWV 良好的可靠性与稳定性。

2)以 GNSS/RS PWV 为参考,验证了 LP PWV 与 LNL PWV 在不同历元的表现情况。新方法获取的 LNL PWV 与 GNSS/RS PWV 均具有很好的一致性,其平均偏差分别为 0.52 mm/0.78 mm,表明新方法获取的 LNL PWV 在不同时序下均具有较好的鲁棒性。此外,以 GNSS PWV 为参考,对新方法获取的水汽空间分布特性进行分析。分析发现 ERA5 PWV 与 PWV 量级和高程相关,而本文方法获取的

LNL PWV 基本不受 PWV 量级与高程影响,具有较好的鲁棒性。

因此,本研究提出的顾及 GNSS 水汽线性和非线性特征的 ERA5 PWV 校正方法,能够很好地提高区域 ERA5 PWV 的精度,且具有良好的鲁棒性,对于获取区域高精度 PWV 数据具有重要意义。

致谢

感谢中国大陆构造环境监测网络提供的 GNSS 观测数据,感谢美国国家气候数据中心提供的 IGRA 数据资料,感谢国家气象中心提供的相关数据。

参考文献

- [1] BEVIS M, BUSINGER S, HERRING T A, et al. GPS meteorology: remote sensing of atmospheric water vapor using the global positioning system[J]. Journal of Geophysical Research: Atmospheres, 1992, 97 (D14): 15787-15801.
- [2] ROSS R J, ELLIOTT W P. Radiosonde-based northern hemisphere tropospheric water vapor trends[J]. Journal of Climate, 2001, 14: 1602-1612.
- [3] ZHANG B, YAO Y, XIN L, et al. Precipitable water vapor fusion: an approach based on spherical cap harmonic analysis and Helmert variance component estimation [J]. Journal of Geodesy, 2019, 93(12): 2605-2620.
- [4] NIELL A E, COSTER A J, SOLHEIM F S, et al. Comparison of measurements of atmospheric wet delay by radiosonde, water vapor radiometer, GPS, and VLBI [J]. Journal of Atmospheric and Oceanic Technology, 2001, 18(6): 830-850.
- [5] ZHU H, CHEN K, CHAI H, et al. Characterizing extreme drought and wetness in Guangdong, China using global navigation satellite system and precipitation data [J]. Satellite Navigation, 2024, 5(1): 1.

- [6] TSIDU G M, BLUMENSTOCK T, HASE F. Observations of precipitable water vapour over complex topography of Ethiopia from ground-based GPS, FTIR, radiosonde and ERA-Interim reanalysis[J]. Atmospheric Measurement Techniques, 2015, 8(8): 3277-3295.
- [7] 张克非, 李浩博, 王晓明, 等. 地基GNSS大气水汽探测遥感研究进展和展望[J]. 测绘学报, 2022, 51(7): 1172-1191.
ZHANG Kefei, LI Haobo, WANG Xiaoming, et al. Recent progresses and future prospectives of ground-based GNSS water vapor sounding[J]. Acta Geodaetica et Cartographica Sinica, 2022, 51(7): 1172-1191 (in Chinese).
- [8] ZHAO Q, DU Z, YAO W, et al. Precipitable water vapor fusion method based on artificial neural network [J]. Advances in Space Research, 2022, 70(1): 85-95.
- [9] ZHANG Y, CAI C, CHEN B, et al. Consistency evaluation of precipitable water vapor derived from ERA5, ERA-Interim, GNSS, and radiosondes over China[J]. Radio Science, 2019, 54(7): 561-571.
- [10] ZHAO Q, YAO Y, YAO W, et al. GNSS-derived PWV and comparison with radiosonde and ECMWF ERA-Interim data over mainland China[J]. Journal of Atmospheric and Solar-Terrestrial Physics, 2019, 182: 85-92.
- [11] 黎峻宇, 李浩杰, 姚宜斌, 等. 利用广义回归神经网络融合天顶对流层湿延迟[J/OL]. 武汉大学学报(信息科学版):1-12(2022-09-25)[2024-07-12]. <https://doi.org/10.13203/j.whugis20220193>.
LI Junyu, LI Haojie, YAO Yibin, et al. Zenith wet delay fusion based on a generalized regression neural network[J/OL]. Geomatics and Information Science of Wuhan University:1-12(2022-09-25)[2024-07-12]. <https://doi.org/10.13203/j.whugis20220193> (in Chinese).
- [12] 刘严萍, 曾昭扬, 王勇. 基于GNSS的中国西南地区MODIS水汽校正研究[J]. 大地测量与地球动力学, 2022, 42(4): 389-392.
LIU Yanping, ZENG Zhaoyang, WANG Yong. MODIS precipitable water vapor correction in southwest China based on GNSS[J]. Journal of Geodesy and Geodynamics, 2022, 42(4): 389-392 (in Chinese).
- [13] 王梦瑶, 张书华, 张文渊, 等. 融合多非线性因素的MODIS PWV神经网络差分校正模型[J/OL]. 测绘学报: 1-13 (2024-04-19) [2024-11-01]. <http://kns.cnki.net/kcms/detail/11.2089.p.20240417.1140.004.html>.
- [14] WANG Mengyao, ZHANG Shubi, ZHANG Wenyuan, et al. MODIS PWV neural network differential correction model integrating multiple nonlinear factors [J/OL]. Acta Geodaetica et Cartographica Sinica: 1-13 (2024-04-19) [2024-11-01]. <http://kns.cnki.net/kcms/detail/11.2089.p.20240417.1140.004.html> (in Chinese).
- [15] 赵庆志, 马智, 姚宜斌, 等. GNSS辅助风云三号卫星MERSI近红外通道的大气可降水量反演方法[J]. 测绘学报, 2024, 53(2): 306-320.
ZHAO Qingzhi, MA Zhi, YAO Yibin, et al. GNSS-assisted FY-3 satellite atmospheric precipitable water retrieval method[J]. Acta Geodaetica et Cartographica Sinica, 2024, 53(2): 306-320 (in Chinese).
- [16] LI X, LONG D. An improvement in accuracy and spatiotemporal continuity of the MODIS precipitable water vapor product based on a data fusion approach [J]. Remote Sensing of Environment, 2020, 248: 111966.
- [17] LU C, ZHANG Y, ZHENG Y, et al. Precipitable water vapor fusion of MODIS and ERA5 based on convolutional neural network [J]. GPS Solutions, 2022, 27(1): 15.
- [18] MA X, YAO Y, ZHANG B, et al. An improved MODIS NIR PWV retrieval algorithm based on an artificial neural network considering the land-cover types[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-12.
- [19] ZHANG B, YAO Y. Precipitable water vapor fusion based on a generalized regression neural network[J]. Journal of Geodesy, 2021, 95(3): 36.
- [20] REN D, WANG Y, WANG G, et al. Fusion of CMONOC and ERA5 PWV products based on back-propagation neural network [J]. Remote Sensing, 2022, 14(15): 3750.
- [21] XIN Z, YU X, LI Q, et al. Spatiotemporal variation in rainfall erosivity on the Chinese Loess Plateau during the period 1956—2008[J]. Regional Environmental Change, 2011, 11(1): 149-159.
- [22] 李澳, 于志静, 王德富, 等. 黄土高原地区水土流失对土地利用和降水变化的响应[J]. 水土保持学报, 2024, 38(4): 95-102+113.
LI Ao, YU Zhijing, WANG Defu, et al. Response of soil and water loss to land use and precipitation changes on the Loess Plateau[J]. Journal of Soil and Water Conservation, 2024, 38(4): 95-102+113 (in Chinese).
- [23] 赵庆志, 杜正, 姚宜斌, 等. 时空加权与再分析资料相结合的GNSS PWV时序填补方法[J]. 测绘学报,

- 2023, 52(10): 1661-1668.
- ZHAO Qingzhi, DU Zheng, YAO Yibin, et al. Combining spatio-temporal weighting with reanalysis data for filing in GNSS PWV time series[J]. *Acta Geodaetica et Cartographica Sinica*, 2023, 52(10): 1661-1668(in Chinese).
- [23] SUN Z, ZHANG B, YAO Y. An ERA5-based model for estimating tropospheric delay and weighted mean temperature over China with improved spatiotemporal resolutions[J]. *Earth and Space Science*, 2019, 6(10): 1926-1941.
- [24] 张晨艳, 何秀凤, 贾东振. 瀑布沟水电站移民新城变形监测与预测分析[J]. *测绘科学技术学报*, 2012, 29(6): 422-426.
- ZHANG Chenyan, HE Xiufeng, JIA Dongzhen. Prediction and deformation of Pubugou hydropower engineering resettlement [J]. *Journal of Geomatics Science and Technology*, 2012, 29(6): 422-426 (in Chinese).
- [25] 黄良珂, 莫智翔, 刘立龙, 等. 顾及时变递减因子的中国大陆地区大气可降水量垂直改正模型[J]. *测绘学报*, 2021, 50(10): 1320-1330.
- HUANG Liangke, MO Zhixiang, LIU Lilong, et al. An empirical model for the vertical correction of precipitable water vapor considering the time varying lapse rate for mainland China[J]. *Acta Geodaetica et Cartographica Sinica*, 2021, 50(10): 1320-1330 (in Chinese).
- [26] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors[J]. *Nature*, 1986, 323(6088): 533-536.
- [27] XU J, LIU Z. A back propagation neural network-based algorithm for retrieving all-weather precipitable water vapor from MODIS NIR measurements[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 1-14.
- [28] ZHAO Q, LIU Y, YAO W, et al. Hourly rainfall forecast model using supervised learning algorithm[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 1-9.
- [29] TAN J, CHEN B, WANG W, et al. Evaluating pre-cipitable water vapor products from Fengyun-4A meteorological satellite using radiosonde, GNSS, and ERA5 data[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 1-12.
- [30] JIANG N, WU Y, LI S, et al. First PWV retrieval using MERSI-LL onboard FY-3E and cross validation with co-platform occultation and ground GNSS[J]. *Geophysical Research Letters*, 2024, 51(8): 108681.
- [31] YAN X, YANG W, DING N, et al. Improving MODIS-IR precipitable water vapor based on the FIDWFT model [J]. *Advances in Space Research*, 2024, 73(10): 4903-4921.
- [32] WANG Y, YANG F, LI P, et al. An optimal calibration method for MODIS precipitable water vapor using GNSS observations[J]. *Atmospheric Research*, 2024, 309: 107591.
- [33] 赵庆志, 刘康, 李祖锋, 等. GNSS 和非实测气象参数的 PWV 反演方法及其精度评估[J]. 武汉大学学报(信息科学版), 2024, 49(3): 453-464.
- ZHAO Qingzhi, LIU Kang, LI Zufeng, et al. PWV inversion method based on GNSS and non-measured meteorological parameters and accuracy evaluation [J]. *Geomatics and Information Science of Wuhan University*, 2024, 49(3): 453-464 (in Chinese).
- [34] MA X, YAO Y, ZHANG B, et al. Retrieval of high spatial resolution precipitable water vapor maps using heterogeneous earth observation data [J]. *Remote Sensing of Environment*, 2022, 278: 113100.
- [35] XIONG Z, SUN X, SANG J, et al. Modify the accuracy of MODIS PWV in China: a performance comparison using random forest, generalized regression neural network and back-propagation neural network [J]. *Remote Sensing*, 2021, 13(11): 2215.
- [36] DHARIA A, ADEI H. Neural network model for rapid forecasting of freeway link travel time[J]. *Engineering Applications of Artificial Intelligence*, 2003, 16 (7): 607-613.
- [37] XIONG Z, ZHANG B, SANG J, et al. Fusing precipitable water vapor data in China at different timescales using an artificial neural network[J]. *Remote Sensing*, 2021, 13(9): 1720.

(编辑:黄利华)